

Requested Patent: JP2002358091A

Title: ;

Abstracted Patent: JP2002358091 ;

Publication Date: 2002-12-13 ;

Inventor(s): ;

Applicant(s): ;

Application Number: JP20010166484 20010601 ;

Priority Number(s): JP20010166484 20010601 ;

IPC Classification: G10L13/06; G06F3/16; G10L13/00; G10L13/08 ;

Equivalents: ;

ABSTRACT:

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2002-358091  
(P2002-358091A)

(43) 公開日 平成14年12月13日 (2002. 12. 13)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テ-マ-ト <sup>*</sup> (参考)
G 1 0 L 13/06		G 0 6 F 3/16	3 3 0 K 5 D 0 4 5
G 0 6 F 3/16	3 3 0	G 1 0 L 5/04	F
G 1 0 L 13/00		3/00	E
13/08			H

審査請求 未請求 請求項の数 6 O L (全 8 頁)

(21) 出願番号 特願2001-166484 (P2001-166484)

(22) 出願日 平成13年6月1日 (2001. 6. 1)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 磯野 敏幸

神奈川県横浜市港北区綱島東四丁目3番1  
号 松下通信工業株式会社内

(72) 発明者 西村 祥文

神奈川県横浜市港北区綱島東四丁目3番1  
号 松下通信工業株式会社内

(74) 代理人 100072604

弁理士 有我 軍一郎

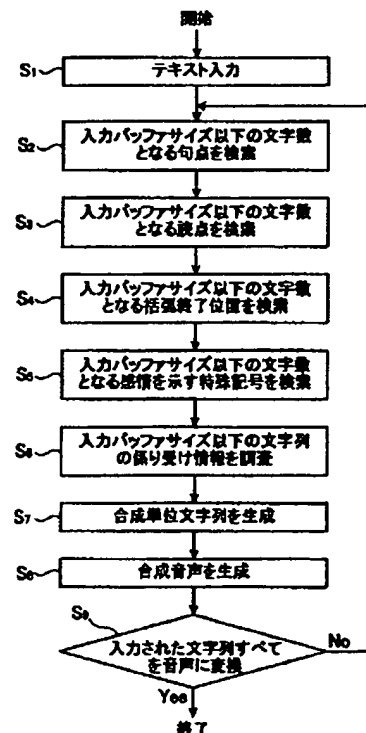
Fターム(参考) 5D045 AA07 AA20

(54) 【発明の名称】 音声合成方法および音声合成装置

(57) 【要約】

【課題】 入力バッファサイズに制限がある場合でも文章を自然に合成音声に変換することができる音声合成方法および音声合成装置を提供する。

【解決手段】 入力したテキストデータの中から入力バッファサイズ以下の文字数となる句読点、括弧の終了位置、感情を示す特殊記号および係り受け情報を検索、調査し (ステップS1～S6)、前記検索、調査結果に基づいて入力したテキストデータから入力バッファサイズ以下の文字列を分離し、合成単位文字列を生成する (ステップS7)。次いで、生成された合成単位文字列に基づいて合成音声を生成し、このように入力されたテキストの文字列の全てが音声に変換されるまで一連の処理を繰り返し、全ての文字列を音声化する (ステップS8、S9)。



## 【特許請求の範囲】

【請求項1】 入力された文字列から入力バッファサイズ以下の文字数となる句読点、括弧の終了位置、感情を示す特殊記号位置および係り受け情報を検出し、前記文字列から入力バッファサイズ以下の長さの合成単位文字列を分離し、分離された合成単位文字列毎に合成音声を生成することを特徴とする音声合成方法。

【請求項2】 前記分離された合成単位文字列の中で感情を示す特殊記号については、感情付与音声素片データベースおよび朗読調音声素片データベースの中から感情付与音声素片データベースを選択して合成音声を生成することを特徴とする請求項1記載の音声合成方法。

【請求項3】 前記分離された合成単位文字列の中で括弧に囲まれた文字列については、口語調音声素片データベースおよび朗読調音声素片データベースの中から口語調音声素片データを選択して合成音声を生成することを特徴とする請求項1または2記載の音声合成方法。

【請求項4】 テキストを入力するテキスト入力手段と、入力バッファサイズ以下の文字数となる句点を検索する句点検索手段と、入力バッファサイズ以下の文字数となる読点を検索する読点検索手段と、入力バッファサイズ以下の文字数となる括弧の終了位置を検索する括弧終了位置検索手段と、入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索する感情を示す特殊記号検索手段と、入力バッファサイズ以下の文字数となる文字列の係り受け情報を調査する係り受け情報調査手段と、前記句点検索手段、前記読点検索手段、前記括弧終了位置検索手段、前記感情を示す特殊記号検索手段、前記係り受け情報調査手段に基づいて入力されたテキストデータ文字列から入力バッファサイズ以下の文字列を分離し、合成単位文字列を生成する合成単位文字列生成手段と、合成音声を生成するための音声素片データベースと、前記合成単位文字列生成手段により生成された文字列を合成音声データに変換する合成音声生成手段と、前記合成音声生成手段により生成された音声信号を音声出力する音声出力手段とを備えることを特徴とする音声合成装置。

【請求項5】 感情が付与された感情付与音声素片データベースと、朗読調音声素片データベースと、前記合成単位文字列生成手段により生成された文字列を検索し、前記文字列の中に感情を示す特殊記号がある場合に、前記感情付与音声素片データベースを選択し、前記文字列の中に感情を示す特殊記号がない場合に、前記朗読調音声素片データベースを選択する音声素片データベース選択手段とを備えることを特徴とする請求項4記載の音声合成装置。

【請求項6】 口語調音声素片データベースと、朗読調音声素片データベースと、前記合成単位文字列生成手段により生成された文字列を検索し、前記文字列の中に括弧に囲まれた文字列がある場合に、前記口語調音声素片

データベースを選択し、前記文字列の中に括弧に囲まれた文字列がない場合に、前記朗読調音声素片データベースを選択する音声素片データベース選択手段とを備えることを特徴とする請求項4または5記載の音声合成装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、入力された文字列に基づいて音声を出力する音声合成方法および音声合成装置に関する。

## 【0002】

【従来の技術】入力された文字列に基づいて音声を合成する方法としては、文字列を感情のある音声や無感情の音声に変化するものがあり、このようなものとして、例えば、特開平7-72900号公報に記載されたようなものがある。

【0003】このものは、文字コードを構文解析し、予め内部に構築されている辞書を参照しながら各単語や文節のアクセントを決定し、かつ同様に内部に構築されている様々な音素を表現できるスペクトルの特徴パラメータの中から該当する単語等を構成するのに最も適切なものを選択し、さらに、規則により文章全体のイントネーションやパワーの変化を決定し、これらに基づいて音声を出力するようにしている。

【0004】この技術にあつては、合成音声の入力バッファサイズに関しては特に制限は設けられてはいなく、実際は冗長的なバッファサイズを設定して、入力された文字列がバッファサイズを越えないようにしていた。

## 【0005】

【発明が解決しようとする課題】しかしながら、このような従来の音声合成方法においては、メモリサイズの都合で入力バッファサイズに制限が生じる場合には、入力バッファサイズ以上の文字数である文字列は合成できないという問題があった。

【0006】本発明は、従来の問題を解決するためになされたもので、入力バッファサイズに制限がある場合でも文章を自然に合成音声に変換することができる音声合成方法および音声合成装置を提供するものである。

## 【0007】

【課題を解決するための手段】本発明の音声合成方法は、入力された文字列から入力バッファサイズ以下の文字数となる句読点、括弧の終了位置、感情を示す特殊記号位置および係り受け情報を検出し、前記文字列から入力バッファサイズ以下の長さの合成単位文字列を分離し、分離された合成単位文字列毎に合成音声を生成するものである。

【0008】このような方法により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、自然な位置で文章を区切ることができ、文意に沿った合成音声を生成することができる。

【0009】また、本発明の音声合成方法は、前記分離された合成単位文字列の中で感情を示す特殊記号については、感情付与音声素片データベースおよび朗読調音声素片データベースの中から感情付与音声素片データベースを選択して合成音声を生成するものである。

【0010】このような方法により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切ることができ、文意に沿った合成音声を生成することができるとともに、感情を示す特殊記号に応じた感情が付与された合成音声を生成することができる。

【0011】また、本発明の音声合成方法は、前記分離された合成単位文字列の中で括弧に囲まれた文字列については、口語調音声素片データベースおよび朗読調音声素片データベースの中から口語調音声素片データを選択して合成音声を生成するものである。

【0012】このような方法により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切ることができ、文意に沿った合成音声を生成することができるとともに、括弧で囲まれた文字列は口語調の合成音声を生成することができる。

【0013】また、本発明の音声合成装置は、テキストを入力するテキスト入力手段と、入力バッファサイズ以下の文字数となる句点を検索する句点検索手段と、入力バッファサイズ以下の文字数となる読点を検索する読点検索手段と、入力バッファサイズ以下の文字数となる括弧の終了位置を検索する括弧終了位置検索手段と、入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索する感情を示す特殊記号検索手段と、入力バッファサイズ以下の文字数となる文字列の係り受け情報を調査する係り受け情報調査手段と、前記句点検索手段、前記読点検索手段、前記括弧終了位置検索手段、前記感情を示す特殊記号検索手段、前記係り受け情報調査手段に基づいて入力されたテキストデータ文字列から入力バッファサイズ以下の文字列を分離し、合成音声単位文字列を生成する合成音声単位文字列生成手段と、合成音声を生成するための音声素片データベースと、前記合成音声単位文字列生成手段により生成された文字列を合成音声データに変換する合成音声生成手段と、前記合成音声生成手段により生成された音声信号を音声出力する音声出力手段とを備えて構成される。

【0014】このような構成により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切ることができ、文意に沿った合成音声を生成することができる音声合成装置を提供することができる。

【0015】また、本発明の音声合成装置は、感情が付与された感情付与音声素片データベースと、朗読調音声素片データベースと、前記合成音声単位文字列生成手段により生成された文字列を検索し、前記文字列の中に感情を

示す特殊記号がある場合に、前記感情付与音声素片データベースを選択し、前記文字列の中に感情を示す特殊記号がない場合に、前記朗読調音声素片データベースを選択する音声素片データベース選択手段とを備えて構成される。

【0016】このような構成により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切ることができ、文意に沿った合成音声を生成することができるとともに、感情を示す特殊記号に応じた感情が付与された合成音声を生成することができる音声合成装置を提供することができる。

【0017】また、本発明の音声合成装置は、口語調音声素片データベースと、朗読調音声素片データベースと、前記合成音声単位文字列生成手段により生成された文字列を検索し、前記文字列の中に括弧に囲まれた文字列がある場合に、前記口語調音声素片データベースを選択し、前記文字列の中に括弧に囲まれた文字列がない場合に、前記朗読調音声素片データベースを選択する音声素片データベース選択手段とを備えて構成される。

【0018】このような構成により、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切ることができ、文意に沿った合成音声を生成することができるとともに、括弧で囲まれた文字列は口語調の合成音声を生成することができる音声合成装置を提供することができる。

【0019】

【発明の実施の形態】以下、本発明の実施形態を図面に基いて説明する。

【0020】図1、図2は本発明に係る音声合成装置および音声合成方法の第1実施形態を示す図である。

【0021】まず、構成を説明する。図1において、本実施形態の音声合成装置は、テキスト入力手段1、句点検索手段2、読点検索手段3、括弧終了位置検索手段4、感情を示す特殊記号検索手段5、係り受け情報調査手段6、合成音声単位文字列生成手段7、音声合成生成手段8、音声素片データベース9および音声出力手段10から構成される。

【0022】テキスト入力手段1にはテキスト文書が入力されるようになっており、句点検索手段2はこの入力されたテキスト文書から入力バッファサイズ以下の文字数となる句点を検索する。また、読点検索手段3は入力バッファサイズ以下の文字数となる読点を検索するようになっており、括弧終了位置検索手段4により入力バッファサイズ以下の文字数となる括弧終了位置を検索する。

【0023】また、特殊記号検索手段5は入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索するようになっており、係り受け情報調査手段6は入力バッファサイズ以下の文字列の係り受け情報を調査するようになっている。

【0024】また、合成単位文字列生成手段7により句点検索手段2、読点検索手段3、括弧終了位置検索手段4、感情を示す特殊記号検索手段5、係り受け情報調査手段6によって得られた情報に基づいて、入力バッファサイズ以下となる文字列をテキストデータ文字列から分離し、合成単位文字列を生成するようになっている。また、合成音声生成手段8は音声素片データベース9を使い合成単位文字列生成手段7により生成された文字列の合成音声を生成するようになっている。

【0025】また、音声出力手段10は合成音声生成手段8で生成された合成音声を出力するようになっている。

【0026】次に、図2に示すフローチャートに基づいて音声合成方法を説明する。

【0027】まず、テキスト文書の入力処理が行なわれると（ステップS1）、句点検索手段2によって入力バッファサイズ以下の文字数となる句点を検索する処理を実行する（ステップS2）。

【0028】次いで、読点検索手段3によって入力バッファサイズ以下の文字数となる読点を検索する処理を実行した後（ステップS3）、括弧終了位置検索手段4によって入力バッファサイズ以下の文字数となる括弧の終了位置を検索する処理を実行する（ステップS4）。次いで、特殊記号検索手段5によって入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索する処理を実行した後（ステップS5）、係り受け情報調査手段6によって入力バッファサイズ以下の文字列の係り受け情報を調査する処理を実行する（ステップS6）。

【0029】次いで、合成単位文字列生成手段7によって合成単位文字列を生成する処理を実行した後（ステップS7）、音声合成生成手段8によって合成音声を生成する処理を実行する（ステップS8）。次いで、入力された文字列すべてが音声に変換されたか否かを判別し（ステップS9）、変換されていた場合には今回の処理を終了し、変換されていない場合には、ステップS2に戻ってステップS2以降の処理を実行して入力されたテキストの文字列の全てが音声に変換されるまで一連の処理を繰り返す、すべての文字列を音声化する。

【0030】このように本実施形態では、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切り、文意に沿った合成音声を生成することができる。

【0031】図3、図4は本発明に係る音声合成方法および音声合成装置の第2実施形態を示す図である。なお、本実施形態では、第1実施形態と同様の構成には、同一番号を付して説明を省略する。

【0032】本実施形態では、第1実施形態と異なる構成は、図3に示すように、音声素片データベース選択手段11、朗読調音声素片データベース12、感情付与音声素片データベース13を有する点である。

【0033】音声素片データベース選択手段11は、合成

単位文字列生成手段7によって生成された合成単位文字列に感情を示す特殊記号が含まれているか否かを判別するようになっており、感情を示す特殊記号が含まれている場合には、感情付与音声素片データベース13を選択し、そうでない場合には朗読調音声素片データベース12を選択するようになっている。

【0034】また、感情付与音声素片データベース13には感情が付与された素片が記憶されており、朗読調音声素片データベース12には感情が付与されていない朗読調の素片が記憶されている。

【0035】次に、図4に基づいて音声合成方法を説明する。

【0036】テキスト文書の入力処理が行なわれると（ステップS11）、句点検索手段2によって入力バッファサイズ以下の文字数となる句点を検索する処理を実行する（ステップS12）。

【0037】次いで、読点検索手段3によって入力バッファサイズ以下の文字数となる読点を検索する処理を実行した後（ステップS13）、括弧終了位置検索手段4によって入力バッファサイズ以下の文字数となる括弧の終了位置を検索する処理を実行する（ステップS14）。次いで、特殊記号検索手段5によって入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索する処理を実行した後（ステップS15）、係り受け情報調査手段7によって入力バッファサイズ以下の文字列の係り受け情報を調査する処理を実行する（ステップS16）。

【0038】次いで、合成単位文字列生成手段7によって合成単位文字列を生成する処理を実行した後（ステップS17）、音声素片データベース選択手段11によって合成単位文字列生成手段7で生成された合成単位文字列に感情を示す特殊記号が含まれているか否かを判別し、感情を示す特殊記号が含まれている場合には、感情付与音声素片データベース13を選択し、そうでない場合には朗読調音声素片データベース12を選択する（ステップS18）。

【0039】次いで、音声合成生成手段8によって選択されたデータベースの合成音声を生成する処理を実行した後（ステップS19）、入力された文字列すべてが音声に変換されたか否かを判別し（ステップS20）、変換されていた場合には今回の処理を終了し、変換されていない場合にはステップS12に戻ってステップS12以降の処理を実行して入力されたテキストの文字列の全てが音声に変換されるまで一連の処理を繰り返す、すべての文字列を音声化する。

【0040】このように本実施形態では、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切り、文意に沿った合成音声を生成できるとともに、感情を示す特殊記号に応じて感情が付与された合成音声を生成することができる。

【0041】図5、図6は本発明に係る音声合成方法お

よび音声合成装置の第3実施形態を示す図である。なお、本実施形態では、第1実施形態と同様の構成には、同一番号を付して説明を省略する。

【0042】本実施形態では、第1実施形態と異なる構成は、図5に示すように、音声素片データベース選択手段21、朗読調音声素片データベース22、口語調音声素片データベース23を有する点である。

【0043】音声素片データベース選択手段11は、合成単位文字列生成手段7によって生成された合成単位文字列の中に括弧に囲まれた文字列があるか否かを判別するようになっており、括弧に囲まれた文字がある場合には、口語調音声素片データベース23を選択し、括弧に囲まれた文字がない場合には、朗読調音声素片データベース22を選択するようになっている。

【0044】また、朗読調音声素片データベース22には感情が付与されていない朗読調の素片が記憶されており、口語調音声素片データベース23には口語調の素片が記憶されている。

【0045】次に、図6に基づいて音声合成方法を説明する。

【0046】テキスト文書の入力処理が行なわれると（ステップS21）、句点検索手段2によって入力バッファサイズ以下の文字数となる句点を検索する処理を実行する（ステップS22）。

【0047】次いで、読点検索手段3によって入力バッファサイズ以下の文字数となる読点を検索する処理を実行した後（ステップS23）、括弧終了位置検索手段4によって入力バッファサイズ以下の文字数となる括弧の終了位置を検索する処理を実行する（ステップS24）。次いで、特殊記号検索手段5によって入力バッファサイズ以下の文字数となる感情を示す特殊記号を検索する処理を実行した後（ステップS25）、係り受け情報調査手段7によって入力バッファサイズ以下の文字列の係り受け情報を調査する処理を実行する（ステップS26）。

【0048】次いで、合成単位文字列生成手段7によって合成単位文字列を生成する処理を実行した後（ステップS27）、音声素片データベース選択手段11によって合成単位文字列生成手段7で生成された合成単位文字列に括弧で囲まれた文字列が含まれているか否かを判別し、括弧で囲まれた文字列が含まれている場合には、口語調音声素片データベース23を選択し、そうでない場合には朗読調音声素片データベース22を選択する（ステップS28）。

【0049】次いで、音声合成生成手段8によって選択されたデータベースの合成音声を生成する処理を実行した後（ステップS29）、入力された文字列すべてが音声に変換されたか否かを判別し（ステップS30）、変換されていた場合には今回の処理を終了し、変換されてい

ない場合にはステップS22に戻ってステップS22以降の処理を実行して入力されたテキストの文字列の全てが音声に変換されるまで一連の処理を繰り返し、すべての文字列を音声化する。

【0050】このように本実施形態では、入力バッファサイズ以上の文字数である文字列が入力された場合にも、文章を自然に区切り、文意に沿った合成音声を生成することができるとともに、括弧に囲まれた文字列は口語調の合成音声を生成することができる。

【0051】

【発明の効果】本発明によれば、入力バッファサイズ以下の文字数となるように入力された文字列を合成単位文字列として分離し、合成単位文字列毎に合成音声に変換することにより、入力バッファサイズの制限がある場合でも文章を自然に合成音声に変換することができるという効果を有する。

【図面の簡単な説明】

【図1】本発明に係る音声合成方法および音声合成装置の第1実施形態を示す図であり、音声合成装置のブロック図

【図2】第1実施形態の音声合成方法のフローチャート

【図3】本発明に係る音声合成方法および音声合成装置の第2実施形態を示す図であり、音声合成装置のブロック図

【図4】第2実施形態の音声合成方法のフローチャート

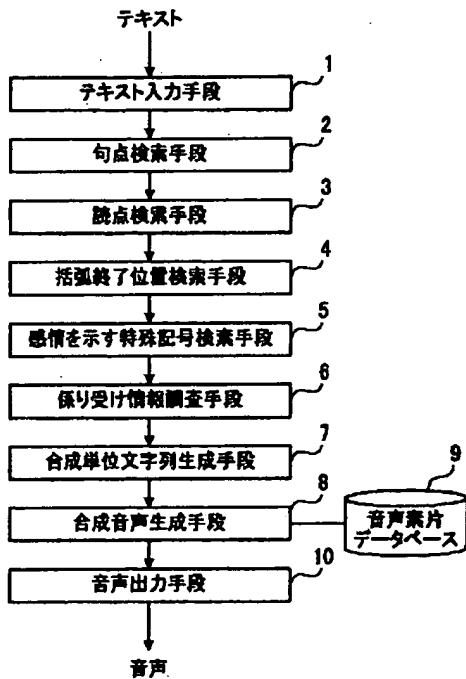
【図5】本発明に係る音声合成方法および音声合成装置の第3実施形態を示す図であり、音声合成装置のブロック図

【図6】第3実施形態の音声合成方法のフローチャート

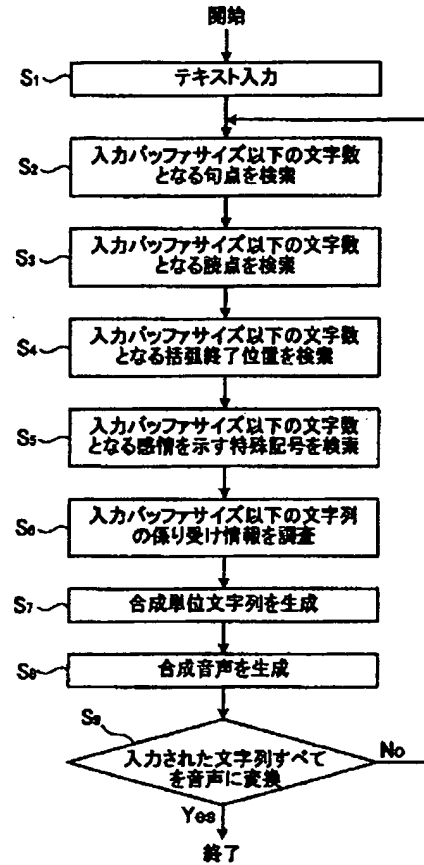
【符号の説明】

- 1 テキスト入力手段
- 2 句点検索手段
- 3 読点検索手段
- 4 括弧終了位置検索手段
- 5 感情を示す特殊記号検索手段
- 6 係り受け情報調査手段
- 7 合成単位文字列生成手段
- 8 音声合成生成手段
- 9 音声素片データベース
- 10 音声出力手段
- 11 データベース選択手段
- 12 朗読調音声素片データベース
- 13 感情付与音声素片データベース
- 21 データベース選択手段
- 22 朗読調音声素片データベース
- 23 口語調音声素片データベース

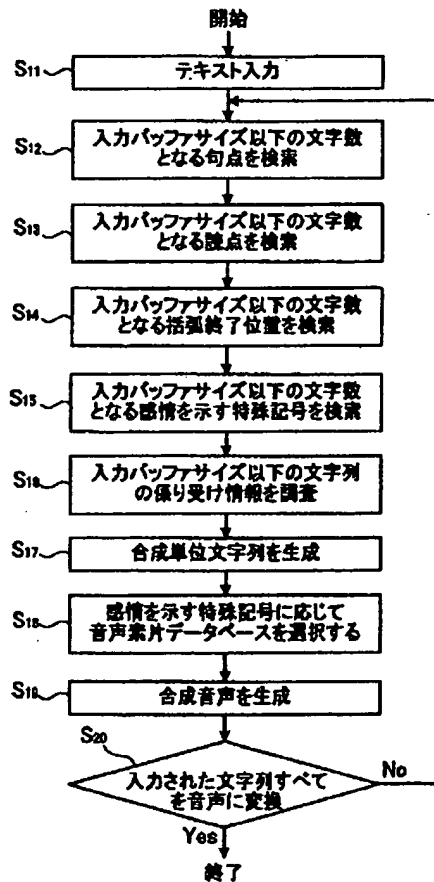
【図1】



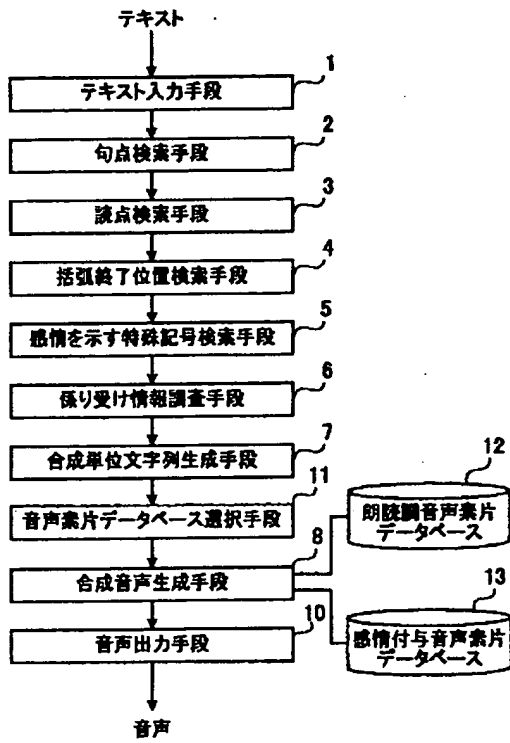
【図2】



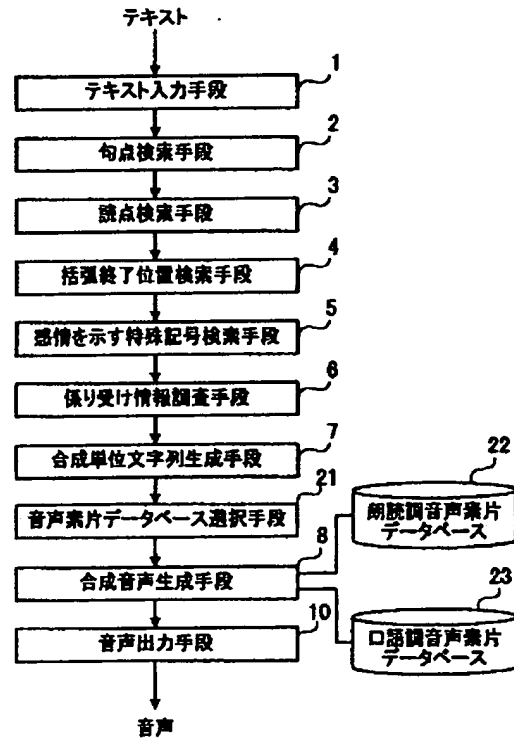
【図4】



【図3】



【図5】





【図6】

